



Deliverable D8.4

Report on eye-movement-based user interface studies

Contract number: **FP7-216529** PinView

Personal Information Navigator Adapting Through Viewing

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under *grant agreement* n° 216529.



Identification sheet

Project ref. no.	FP7-216529
Project acronym	PinView
Status and version	Final, Revision: 1.00
Contractual date of delivery	31.12.2009
Actual date of delivery	28.12.2009
Deliverable number	D8.4
Deliverable title	Report on eye-movement-based user interface studies
Nature	report
Dissemination level	Report: PU – Public
WP contributing to the deliverable	WP8 Personal Information Navigator
Task contributing to the deliverable	Task 8.4 Study of proactive eye-movement-based user interface
WP responsible	Teknillinen korkeakoulu
Task responsible	Teknillinen korkeakoulu
Editor	Arto Klami, <arto.klami@tkk.fi>
Editor address	P.O.BOX 5400, FI-02015 TKK, Finland
Authors in alphabetical order	Samuel Kaski, Arto Klami, László Kozma
EC Project Officer	Pierre-Paul Sondag
Keywords	gaze-based interface, image retrieval, implicit feedback
Abstract	This report presents the output of Task 8.4, study of proactive eye-movement-based user interface. A novel zoomable user interface was developed for purely eye-movement based image retrieval, inspired by the predictive eye-typer Dasher. The display shows images as co-centric circles, and implicit relevance feedback inferred from gaze alone is used for retrieving more relevant images when the user zooms in.

List of annexes

kozma09.pdf – Publication ‘‘GaZIR: Gaze-based Zooming Interface for Image Retrieval’’

Contents

1 Overview	4
2 Interface	4
3 Retrieval system and results	4
4 Conclusions	5

1 Overview

This deliverable constitutes the output of Task 8.4 *Study of proactive eye-movement-based user interface* of the *Personal Information Navigator Adapting Through Viewing*, PinView, project, funded by the European Community's Seventh Framework Programme under Grant Agreement n° 216529.

The interface for showing the images is a crucial element in a retrieval engine. The data collection campaign of Task 8.3 and the first studies on inferring image relevance from gaze (Task 2.1, Task 5.1) used the safe version of showing the retrieved images as an array. The purpose of this task was to study alternative displays, both for improved accuracy in relevance prediction and for more intuitive and pleasing user experience.

The main contribution of this task is a novel zoomable interface for image browsing and retrieval, presented in the publication *GaZIR: Gaze-based Zooming Interface for Image Retrieval* published in the 11th International Conference on Multimodal Interfaces and the 6th Workshop Machine Learning in Multimodal Interfaces, organised in Boston, MA, on November 2009. The publication is included as an annex of this deliverable [3]. A survey on alternative interfaces and more extensive empirical results can be found in the MSc thesis of László Kozma, titled *A Proactive Interface for Image Retrieval* [2].

The remainder of this report includes a brief overview of the GaZIR zoomable interface. The details are presented in [3] which is a part of this deliverable.

2 Interface

The browsing interface of GaZIR is designed to collect maximal amount of information from gaze while still being a natural interface for browsing the image collection. The basic interface is illustrated in Figure 1, showing three concentric rings of images. The outermost ring contains the first ten images shown to the user, the second ring shows images retrieved given the relevance feedback collected from the outermost ring, and the innermost ring takes into account feedback from the two previous rings. The user can zoom the interface inwards and outwards. When zooming inwards the system retrieves another set of images, using all the previous images and their estimated relevancies as feedback, and eventually the older rings will disappear from the display.

The concentric rings of images were chosen instead of the standard grid-based thumbnail display used in earlier studies [1, 4], in order to avoid imposing gaze trajectories based on the structure of the display instead of the content. On a standard grid the users are likely to go through the images in a row-by-row manner, considerably lowering the amount of relevance information the eye movements contain. Completely random placement of images would break this pattern optimally, but a user is likely to find such an interface unpleasant to use. A circle of images provides a compromise between these two goals. It does not lead to scanning patterns as strongly fixed as a grid would, allowing image content to play bigger role in determining where to look, yet it is sufficiently close to standard user interfaces to feel intuitive.

3 Retrieval system and results

GaZIR couples the interface with PicSOM retrieval engine. The publication [3] includes the first empirical experiments where the user is using an integrated retrieval system using gaze-based implicit feedback. The results, reported in detail in the accompanying publication, are promising. As shown in Figure 2, gaze-based feedback is always more accurate than the random baseline, whereas the relative performance compared to click-based feedback on the same interface depends heavily on the search task. For some tasks the accuracy is still quite

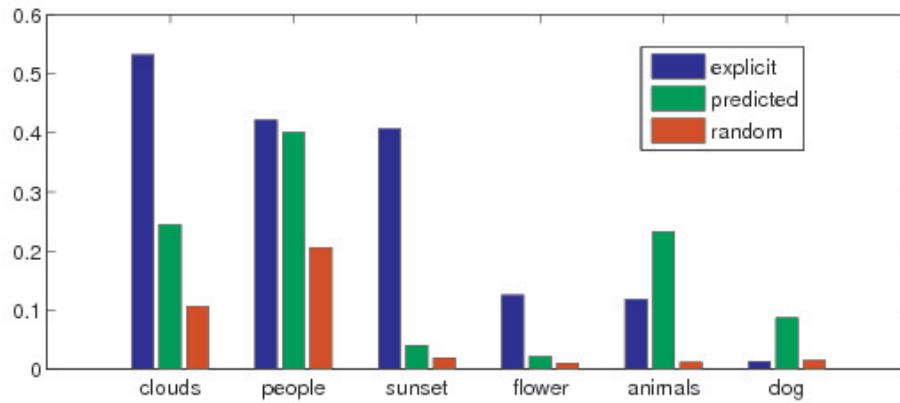


Figure 2: Retrieval performance in real user experiments. The bars indicate the proportion of relevant images shown during the search in six different search tasks for three different feedback methods. Explicit denotes the standard point-and-click feedback, predicted means implicit feedback inferred from gaze, and random is the baseline of providing random feedback. In all cases both actual feedback types outperform the baseline, but the relative performance of explicit and implicit feedback depends on the search task.

References

- [1] Arto Klami, Teofilo de Campos, Kitsuchart Pasupa, Craig Saunders, and Samuel Kaski. Prediction of relevance of an image from a scan pattern. Technical report, PinView, December 31 2008. Deliverable 2.1, available at <http://www.pinview.eu/>.
- [2] László Kozma. A proactive interface for image retrieval. Master's thesis, Helsinki University of Technology, 2009.
- [3] László Kozma, Arto Klami, and Samuel Kaski. GaZIR: Gaze-based zooming interface for image retrieval. In *Proc. ICMI-MLMI 2009, The Eleventh International Conference on Multimodal Interfaces and The Sixth Workshop on Machine Learning for Multimodal Interaction*, pages 305–312, New York, NY, USA, 2009. ACM.
- [4] Craig Saunders and Arto Klami. Database of eye-movement recordings. Technical report, PinView European Community project FP7-216529, July 01 2008. Deliverable D8.3, available at <http://www.pinview.eu/>.